MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Lie Product of Matrix Functions and Its Application to Reachability Analysis for Dynamical Systems with Matrix-valued States

Tomotake SASAKI, Shinji HARA
and Koji TSUMURA

(Communicated by Kazuo MUROTA)

METR 2009–08 March 2009

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

Lie Product of Matrix Functions and Its Application to Reachability Analysis for Dynamical Systems with Matrix-valued States

Tomotake SASAKI, Shinji HARA and Koji TSUMURA

Department of Information Physics and Computing
Graduate School of Information Science and Technology
The University of Tokyo
{Tomotake_Sasaki, Shinji_Hara, Koji_Tsumura}@ipc.i.u-tokyo.ac.jp

March 31st, 2009

Abstract

Time evolutions of some important control targets such as quantum sate or attitude of a rigid body are described by matrix-valued differential equations. In this report, we introduce "Lie product of matrix functions" as a new tool for differential geometric approach to analysis of such systems. A direct and efficient way to calculate the Lie product of matrix functions is provided, which enables us to study important properties of the systems such as reachability or controllability with clear perspective. The effectiveness of the proposed method is confirmed by analysis examples of quantum state control and attitude control of a rigid body.

1 Introduction

In typical cases, time evolution of a control target is described by a differential equation whose variable is a real number vector $\boldsymbol{x}(t) \in \mathbb{R}^d$. However, some important quantities such as quantum states of quantum mechanical systems and attitude of rigid body dynamics are described by matrices. In such cases, time evolution of a control target is described by a matrix-valued differential equation. For example, the time evolution of a class of quantum mechanical systems is described by the following differential equation (Lindblad master equation) whose state variable is an $n \times n$ Hermitian matrix

$X(t)$ [5]:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \sum_{r=1}^{q} (L_r X L_r^* - \frac{1}{2} L_r^* L_r X - \frac{1}{2} X L_r^* L_r) - \mathrm{i}[H, X],$$
$$X(0) = X_0, \tag{1}$$

where H is an $n \times n$ Hermitian matrix (Hamiltonian) and $L_r, r = 1, ..., q$ are $n \times n$ complex matrices. The initial value X_0 is positive semidefinite and unital-trace, or $\mathrm{Tr}\,[X_0] = 1$.

In this report we consider the following differential equation whose state variable is a matrix $X(t) \in \mathbf{X}_n$, which is a generalization of the equation (1):

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mathcal{F}(X) + \sum_{i=1}^{m} \mathcal{G}_i(X) u_i,$$
$$X(0) = X_0 \in \mathbf{X}_n, \tag{2}$$

where $\mathbf{X}_n$ is the set of all $n \times n$ Hermitian matrices denoted by $\mathbf{H}_n$ or the set of all $n \times n$ real matrices denoted by $\mathbf{M}_n(\mathbb{R})$. $\mathcal{F}, \mathcal{G}_i, i = 1, ..., m$ are nonlinear C^∞ functions from $\mathbf{X}_n$ to $\mathbf{X}_n$ which are called C^∞ matrix functions in this report, and $u_i(t) \in \mathbb{R}, i = 1, ..., m$ represents the control inputs. We consider not only linear functions but also nonlinear functions in this report, because we need to treat them in a class of quantum control problems [9]. Furthermore, the equation (2) unifies the equations defined on $\mathbf{H}_n$ and $\mathbf{M}_n(\mathbb{R})$ [1]. This is possible because $\mathbf{H}_n$ and $\mathbf{M}_n(\mathbb{R})$ can be treated essentially in a same manner as seen in the next section.

Differential geometric approach to nonlinear control is useful to analyze important properties of nonlinear systems such as reachability and controllability [8], and it has been used to study controlled quantum systems [1, 11, 10]. However, in these works, differential geometric approach has been applied after representing the equations in a vector-valued form by transforming the variable into a real number vector. A more direct method without such variable transformation is naturally desirable.

In this report, we define "Lie product of matrix functions" as a new tool for differential geometric approach to analyze the systems described by equation (2), and provide an efficient calculation rule which does not require any variable transformation nor elementwise calculation with coordinate expression. Application of the Lie product of matrix functions with the calculation rule simplifies reachability or controllability analysis for the systems described by equation (2). It is expected to be an effective analysis method especially for controlled quantum systems. The effectiveness of the proposed method is verified by examples.

[1] We show an example of equations defined on $\mathbf{M}_n(\mathbb{R})$ in Section 4.

Notation: In this report, we use the following notation. i: imaginary unit. $X^\top$: transpose of a matrix X. X^*: Hermitian conjugate of a matrix X. $[X, Y] = XY - YX$: commutator of two matrices X and Y. $\mathrm{Tr}\,[X]$: trace of a matrix X. E^n_{kl}: $n \times n$ matrix whose (k, l) component is 1 and all the other components are 0.

2 Mathematical Preliminaries

This section is devoted to mathematical preliminaries to define Lie product of matrix functions and to provide an efficient rule for the calculation.

2.1 Matrix Functions and Vector Fields

In this subsection, we clarify the relationship between matrix functions and vector fields. This is the basis of the later discussions.

First, we see some properties of $\mathbf{X}_n$ $(= \mathbf{H}_n, \mathbf{M}_n(\mathbb{R}))$. $\mathbf{X}_n$ is an $N(:= n^2)$ dimensional real Hilbert space. Here for this space we introduce the Frobenius inner product

$$(X, Y)_{\mathrm{F}} := \mathrm{Tr}\,[X^* Y] \tag{3}$$

and the Frobenius norm

$$\|X\|_{\mathrm{F}} := \sqrt{(X, X)_{\mathrm{F}}} \tag{4}$$

for $X, Y \in \mathbf{X}_n$. The set of N Hermitian matrices $\{E^n_{kk}\}_{1 \le k \le n} \cup \{1/\sqrt{2}(E^n_{kl} + E^n_{lk}), \mathrm{i}/\sqrt{2}(-E^n_{kl} + E^n_{lk})\}_{1 \le k < l \le n}$, i.e.,

$$\begin{bmatrix} 1 & & & \mathbf{0} \\ & 0 & & \\ & & \ddots & \\ \mathbf{0} & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & \mathbf{0} \\ & 1 & & \\ & & 0 & \\ & & & \ddots \\ \mathbf{0} & & & 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 & & & \mathbf{0} \\ & \ddots & & \\ & & 0 & \\ \mathbf{0} & & & 1 \end{bmatrix},$$
$$\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & & \mathbf{0} \\ 1 & 0 & & \\ & & \ddots & \\ \mathbf{0} & & & 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -\mathrm{i} & & \mathbf{0} \\ \mathrm{i} & 0 & & \\ & & \ddots & \\ \mathbf{0} & & & 0 \end{bmatrix}, \cdots,$$
$$\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & & & \mathbf{0} \\ & \ddots & & \\ & & 0 & 1 \\ \mathbf{0} & & 1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & & & \mathbf{0} \\ & \ddots & & \\ & & 0 & -\mathrm{i} \\ \mathbf{0} & & \mathrm{i} & 0 \end{bmatrix} \tag{5}$$

is an orthonormal basis of $\mathbf{H}_n$, while the set of N real matrices $\{E^n_{kl}\}_{1 \le k,l \le n}$ is an orthonormal basis of $\mathbf{M}_n(\mathbb{R})$.

According to these choices of orthonormal basises, we can parameterize $X \in \mathbf{H}_n$ as

$$X = \begin{bmatrix} x^1 & \frac{1}{\sqrt{2}}x^{n+1} - \mathrm{i}\frac{1}{\sqrt{2}}x^{n+2} & \cdots \\ \frac{1}{\sqrt{2}}x^{n+1} + \mathrm{i}\frac{1}{\sqrt{2}}x^{n+2} & x^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \tag{6}$$

and $X \in \mathbf{M}_n(\mathbb{R})$ as

$$X = \begin{bmatrix} x^1 & x^2 & \cdots & x^n \\ x^{n+1} & x^{n+2} & \cdots & x^{2n} \\ \vdots & & & \vdots \\ x^{N-n} & x^{N-n+1} & \cdots & x^N \end{bmatrix}, \tag{7}$$

where $x^j \in \mathbb{R}, j = 1, ..., N$. This is the most fundamental characterization of $\mathbf{X}_n$.

We next characterize $\mathbf{X}_n$ from the differential geometric viewpoint and associate matrix functions with vector fields. Refer the Chapter 2 of [8] for the definitions of basic terms such as manifold, tangent vector, tangent space, and vector field.

Since $\mathbf{X}_n$ is a linear space, $\mathbf{X}_n$ can be regarded as an N dimensional differentiable manifold in a natural way [3]. Here we define $\varphi : \mathbf{X}_n \to \mathbb{R}^N$ as

$$\varphi(X) = (\varphi^1(X), ..., \varphi^N(X)) := (x^1, ..., x^N) \tag{8}$$

based on (6) in the case $\mathbf{X}_n = \mathbf{H}_n$ and based on (7) in the case of $\mathbf{X}_n = \mathbf{M}_n(\mathbb{R})$. Note that φ is a one-to-one linear mapping (linear isomorphism), and each $\varphi^j : \mathbf{X}_n \to \mathbb{R}, j = 1, ..., N$ is also linear. The pair $(\mathbf{X}_n, \varphi)$ is a coordinate neighborhood which covers the whole space $\mathbf{X}_n$. We fix the coordinate neighborhood in the following discussions.

Let $T_X\mathbf{X}_n$ be the tangent space of $\mathbf{X}_n$ at a point $X \in \mathbf{X}_n$. We denote the natural basis of $T_X\mathbf{X}_n$ with respect to $(\mathbf{X}_n, \varphi)$ by

$$\left.\frac{\partial}{\partial x^1}\right|_X, ..., \left.\frac{\partial}{\partial x^N}\right|_X. \tag{9}$$

$T_X\mathbf{X}_n$ can be identified with $\mathbf{X}_n$ because $\mathbf{X}_n$ is a linear space [3]. A tangent vector

$$v = \sum_{j=1}^{N} v^j \left.\frac{\partial}{\partial x^j}\right|_X \in T_X\mathbf{X}_n, \tag{10}$$

where $v^j \in \mathbb{R}, j = 1, ..., N$, corresponds to the Hermitian matrix

$$V = \begin{bmatrix} v^1 & \frac{1}{\sqrt{2}}v^{n+1} - \mathrm{i}\frac{1}{\sqrt{2}}v^{n+2} & \cdots \\ \frac{1}{\sqrt{2}}v^{n+1} + \mathrm{i}\frac{1}{\sqrt{2}}v^{n+2} & v^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

and real matrix

$$V = \begin{bmatrix} v^1 & v^2 & \cdots & v^n \\ v^{n+1} & v^{n+2} & \cdots & v^{2n} \\ \vdots & & & \vdots \\ v^{N-n} & v^{N-n+1} & \cdots & v^N \end{bmatrix} \tag{11}$$

in the cases of $\mathbf{X}_n = \mathbf{H}_n$ and $\mathbf{X}_n = \mathbf{M}_n(\mathbb{R})$, respectively.

This correspondence $v \mapsto V$ is obviously one-to-one and linear. We denote this one-to-one linear mapping (linear isomorphism) by

$$\psi_X : T_X \mathbf{X}_n \to \mathbf{X}_n. \tag{12}$$

Let $V^\infty(\mathbf{X}_n)$ denote the set of all C^∞ vector fields defined on $\mathbf{X}_n$ and $C^\infty(\mathbf{X}_n, \mathbf{X}_n)$ represents the set of all C^∞ functions from $\mathbf{X}_n$ to $\mathbf{X}_n$ (C^∞ matrix functions). $V^\infty(\mathbf{X}_n)$ and $C^\infty(\mathbf{X}_n, \mathbf{X}_n)$ have natural linear space structures. Based on the identification of $T_X \mathbf{X}_n$ with $\mathbf{X}_n$ by ψ_X, a C^∞ vector field f that assigns the tangent vector

$$f(X) = \sum_{j=1}^{N} f^j(X) \left.\frac{\partial}{\partial x^j}\right|_X \in T_X \mathbf{X}_n \tag{13}$$

to each point $X \in \mathbf{X}_n$ can be identified with the C^∞ matrix function $\mathcal{F}$ that satisfies $\mathcal{F}(X) = \psi_X(f(X))$ for all $X \in \mathbf{X}_n$. Here $f^j, j = 1, ..., N$ are C^∞ functions from $\mathbf{X}_n$ to $\mathbb{R}$. The matrix function satisfying the above condition has the forms

$$\mathcal{F}(X) = \begin{bmatrix} f^1(X) & \frac{1}{\sqrt{2}} f^{n+1}(X) - \mathrm{i}\frac{1}{\sqrt{2}} f^{n+2}(X) & \cdots \\ \frac{1}{\sqrt{2}} f^{n+1}(X) + \mathrm{i}\frac{1}{\sqrt{2}} f^{n+2}(X) & f^2(X) & \cdots \\ \vdots & & \ddots \end{bmatrix} \tag{14}$$

and

$$\mathcal{F}(X) = \begin{bmatrix} f^1(X) & f^2(X) & \cdots & f^n(X) \\ f^{n+1}(X) & f^{n+2}(X) & \cdots & f^{2n}(X) \\ \vdots & & & \vdots \\ f^{N-n}(X) & f^{N-n+1}(X) & \cdots & f^N(X) \end{bmatrix} \tag{15}$$

in the cases of $\mathbf{X}_n = \mathbf{H}_n$ and $\mathbf{X}_n = \mathbf{M}_n(\mathbb{R})$, respectively.

It is obvious that the correspondence $f \mapsto F$ defined as above is one-to-one and linear. We denote this one-to-one linear mapping by

$$\psi : V^\infty(\mathbf{X}_n) \to C^\infty(\mathbf{X}_n, \mathbf{X}_n). \tag{16}$$

2.2 Gâteaux Differential

The Gâteaux differential is a directional derivative defined for a function whose range is a normed linear space [7]. Definition 2.1 below provides the definition of the Gâteaux differential of a function from $\mathbf{X}_n$ to $\mathbf{X}_n$.

Definition 2.1 (Gâteaux differential of a matrix function) *Let $\mathcal{F}$ be a function from $\mathbf{X}_n$ to $\mathbf{X}_n$ and let X and Y be elements of $\mathbf{X}_n$. If the limit*

$$\frac{\partial \mathcal{F}}{\partial X}(Y) := \lim_{\substack{\epsilon \in \mathbb{R} \\ \epsilon \rightarrow 0}} \frac{\mathcal{F}(X + \epsilon Y) - \mathcal{F}(X)}{\epsilon} \tag{17}$$

exists, it is called the Gâteaux differential of $\mathcal{F}$ at X with the increment Y.

In the later discussions, the Gâteaux differential makes it possible to analyze the reachability or controllability for the systems described by equation (2) with simple calculations.

Here we show several examples of matrix functions.

Example 2.1 *For any $n \times n$ Hermitian matrix A and $n \times n$ complex matrix C, the following functions $\mathcal{F}_j (j = 1, 2, 3, 4)$ are C^∞ functions from $\mathbf{H}_n$ to $\mathbf{H}_n$:*

$$\mathcal{F}_1[A](X) := -\mathrm{i}[A, X], \tag{18}$$

$$\mathcal{F}_2[C](X) := CXC^*, \tag{19}$$

$$\mathcal{F}_3[C](X) := CX + XC^*, \tag{20}$$

$$\mathcal{F}_4[C](X) := \mathrm{Tr}\left[(C + C^*)X\right] X. \tag{21}$$

Note especially that $\mathcal{F}_4[C]$ is a nonlinear function. Using these symbols, components of the Lindblad master equation (1) can be written as $\mathcal{F}_1[H], \mathcal{F}_2[L_r]$, $\mathcal{F}_3[L_r^* L_r]$. $\mathcal{F}_4$ is a function which appears in a class of quantum control problems [9].

Example 2.2 *For any $n \times n$ real matrix K,*

$$\mathcal{F}_5[K](X) := KX \tag{22}$$

is a C^∞ function from $\mathbf{M}_n(\mathbb{R})$ to $\mathbf{M}_n(\mathbb{R})$. Functions of this type appear in attitude control problem of a rigid body (see Section 4).

We have the following results.

Proposition 2.1 *The following relation holds for the* C^∞ *matrix functions* $\mathcal{F}_j (j = 1 \sim 5)$ *defined by (21) and (22):*

$$\frac{\partial \mathcal{F}_1[A]}{\partial X}(Y) = -\mathrm{i}[A, Y], \tag{23}$$

$$\frac{\partial \mathcal{F}_2[C]}{\partial X}(Y) = CYC^*, \tag{24}$$

$$\frac{\partial \mathcal{F}_3[C]}{\partial X}(Y) = CY + YC^*, \tag{25}$$

$$\frac{\partial \mathcal{F}_4[C]}{\partial X}(Y) = \mathrm{Tr}\,[(C + C^*)Y]\,X + \mathrm{Tr}\,[(C + C^*)X]\,Y, \tag{26}$$

$$\frac{\partial \mathcal{F}_5[K]}{\partial X}(Y) = KY, \tag{27}$$

where X, Y *are the elements of respective domain of each function.*

Proof *Taking limit according to (17), we have*

$$\frac{\partial \mathcal{F}_5[K]}{\partial X}(Y) = \lim_{\substack{\epsilon \in \mathbb{R} \\ \epsilon \to 0}} \frac{K(X + \epsilon Y) - KX}{\epsilon} = KY. \tag{28}$$

The other equations (23) $\sim$ *(26) can be obtained in the same way.*

3 Lie Product of Matrix Functions with an Efficient Calculation Rule

In this section, we define Lie product of matrix functions and explain how it works for reachability and controllability analysis for the systems described by equation (2). We then show that the Lie product of matrix functions can be calculated by Gâteaux differential. This provides us an efficient calculation rule for Lie product of matrix functions.

The Lie product (Lie bracket) of vector fields plays an important role in geometric approach to nonlinear control. The definition of the Lie product (Lie bracket) $[f, g]_{\mathrm{L}}$ of C^∞ vector fields f, g on $\mathbf{X}_n$ is given as follows [2].

Definition 3.1 *Let* f, g *be* C^∞ *vector fields on* $\mathbf{X}_n$. *The* Lie product (Lie bracket) $[f, g]_{\mathrm{L}}$ of f and g *is the* C^∞ *vector field that assigns the tangent vector*

$$[f, g]_{\mathrm{L}}(X) = \sum_{k=1}^{N} \left(\sum_{j=1}^{N} \left(\frac{\partial g^k}{\partial x^j}(X) f^j(X) - \frac{\partial f^k}{\partial x^j}(X) g^j(X) \right) \right) \left. \frac{\partial}{\partial x^k} \right|_X \in T_X \mathbf{X}_n \tag{29}$$

[2]See [8] for the coordinate free definition of the Lie product of vector fields. Although the definition given here is in different form, it is equivalent to the original one in the problem setting in this report.

to each point $X \in \mathbf{X}_n$. Here $(\partial g^k/\partial x^j)(X)$ is an abbreviated notation for

$$\frac{\partial(g^k \circ \varphi^{-1})}{\partial x^j}(\varphi^1(X), ..., \varphi^N(X)), \tag{30}$$

that is, the partial derivative of $g^k \circ \varphi^{-1} : \mathbb{R}^N \to \mathbb{R}$ with respect to its j-th component at the point $(\varphi^1(X), ..., \varphi^N(X))$. The same holds for $(\partial f^k/\partial x^j)(X)$.

For $\mathcal{F}, \mathcal{G}_i, i = 1, ..., m$ in (2), define $f, g_i, i = 1, ..., m$ as $f := \psi^{-1}(\mathcal{F}), g_i := \psi^{-1}(\mathcal{G}_i), i = 1, ..., m$. The dimension of the linear subspace of $T_X\mathbf{X}_n$ spanned by $f(X), g_i(X), i = 1, ..., m$ and

$$[f, g_1]_{\mathrm{L}}(X), ..., [f, g_m]_{\mathrm{L}}(X), [f, [f, g_1]_{\mathrm{L}}]_{\mathrm{L}}(X), ... \tag{31}$$

expresses the degree of freedom of local change of the system at state $X \in \mathbf{X}_n$ [8]. The basis for reachability and controllability analysis is the calculation of (31). However, the elementwise calculation of the Lie product of vector fields according to (29) is troublesome, especially in high or general dimensional case. It prevents efficient analysis with clear perspective. We here develop a method to overcome this difficulty, which is more direct one compared with the vector representation method used in [1, 11, 10].

We see in Subsection 2.1 that $V^\infty(\mathbf{X}_n)$ can be identified with $C^\infty(\mathbf{X}_n, \mathbf{X}_n)$. Based on this identification, we define the Lie product of matrix functions as follows.

Definition 3.2 (Lie product of matrix functions) *Let $\mathcal{F}, \mathcal{G}$ be C^∞ functions from $\mathbf{X}_n$ to $\mathbf{X}_n$. The* Lie product $[\mathcal{F}, \mathcal{G}]_{\mathcal{L}}$ of $\mathcal{F}$ and $\mathcal{G}$ *is the C^∞ function from $\mathbf{X}_n$ to $\mathbf{X}_n$ defined as*

$$[\mathcal{F}, \mathcal{G}]_{\mathcal{L}} := \psi\left([\psi^{-1}(\mathcal{F}), \psi^{-1}(\mathcal{G})]_{\mathrm{L}}\right). \tag{32}$$

Note that the Lie product of matrix functions is defined so that the following diagram becomes commutative:

$$\begin{array}{ccc} V^\infty(\mathbf{X}_n) \times V^\infty(\mathbf{X}_n) & \xrightarrow{\psi\times\psi} & C^\infty(\mathbf{X}_n, \mathbf{X}_n) \times C^\infty(\mathbf{X}_n, \mathbf{X}_n) \\ {\scriptstyle [\cdot,\cdot]_{\mathrm{L}}}\big\downarrow & & \big\downarrow{\scriptstyle [\cdot,\cdot]_{\mathcal{L}}} \\ V^\infty(\mathbf{X}_n) & \xrightarrow{\psi} & C^\infty(\mathbf{X}_n, \mathbf{X}_n). \end{array} \tag{33}$$

In other words, we "transplant" the operation $[\cdot, \cdot]_{\mathrm{L}}$ to the world of the matrix functions using ψ.

By Definition 3.2 and the identification of $T_X\mathbf{X}_n$ with $\mathbf{X}_n$, (31) can be identified with

$$[\mathcal{F}, \mathcal{G}_1]_{\mathcal{L}}(X), ..., [\mathcal{F}, \mathcal{G}_m]_{\mathcal{L}}(X), [\mathcal{F}, [\mathcal{F}, \mathcal{G}_1]_{\mathcal{L}}]_{\mathcal{L}}(X), ..., \tag{34}$$

and thus the linear subspace of $T_X\mathbf{X}_n$ spanned by $f(X), g_i(X), i = 1, ..., m$ and (31) can be identified with

$$\mathscr{C}(X) := \text{span}\{\mathcal{F}(X), \mathcal{G}_1(X), ..., \mathcal{G}_m(X), [\mathcal{F}, \mathcal{G}_1]_{\mathcal{L}}(X), ..., [\mathcal{F}, [\mathcal{F}, \mathcal{G}_1]_{\mathcal{L}}]_{\mathcal{L}}(X), ...\}. \tag{35}$$

Consequently, the calculation of Lie product of vector fields for reachability and controllability analysis can be converted to the calculation of Lie product of matrix functions. The mapping $\mathscr{C}$ which assigns a linear subspace of $\mathbf{X}_n$ to each X defined as above can be identified with the *reachability (accessibility) distribution* [8].

The main result of this report is as follows.

Theorem 3.1 *Let $\mathcal{F}$ and $\mathcal{G}$ be C^∞ functions form $\mathbf{X}_n$ to $\mathbf{X}_n$. The following relation holds for the Lie product of $\mathcal{F}$ and $\mathcal{G}$:*

$$[\mathcal{F}, \mathcal{G}]_{\mathcal{L}}(X) = \frac{\partial \mathcal{G}}{\partial X}(\mathcal{F}(X)) - \frac{\partial \mathcal{F}}{\partial X}(\mathcal{G}(X)), \tag{36}$$

where $(\partial\mathcal{G}/\partial X)(\mathcal{F}(X))$ and $(\partial\mathcal{F}/\partial X)(\mathcal{G}(X))$ are Gâteaux differentials of matrix functions given in Definition 2.1.

Proof *Let $f := \psi^{-1}(\mathcal{F})$ and $g := \psi^{-1}(\mathcal{G})$. Regarding the right hand side of (29), the following relation holds:*

$$\sum_{j=1}^{N} \frac{\partial g^k}{\partial x^j}(X) f^j(X) = \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \frac{g^k(X + \epsilon\mathcal{F}(X)) - g^k(X)}{\epsilon}. \tag{37}$$

See Appendix A for the derivation. We can rewrite the term $\sum_{j=1}^{N} \frac{\partial f^k}{\partial x^j}(X) g^j(X)$ in the same way, and we have

$$\begin{aligned}
&[\mathcal{F}, \mathcal{G}]_{\mathcal{L}}(X) \\
&= \psi_X\left([\psi^{-1}(\mathcal{F}), \psi^{-1}(\mathcal{G})](X)\right) \\
&= \psi_X\left(\sum_{k=1}^{N}\left(\lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \frac{g^k(X + \epsilon\mathcal{F}(X)) - g^k(X)}{\epsilon}\right.\right. \\
&\qquad \left.\left. - \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \frac{f^k(X + \epsilon\mathcal{G}(X)) - f^k(X)}{\epsilon}\right) \left.\frac{\partial}{\partial x^k}\right|_X\right) \\
&= \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \psi_X\left(\sum_{k=1}^{N}\left(\frac{g^k(X + \epsilon\mathcal{F}(X)) - g^k(X)}{\epsilon}\right) \left.\frac{\partial}{\partial x^k}\right|_X\right) \\
&\qquad - \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \psi_X\left(\sum_{k=1}^{N}\left(\frac{f^k(X + \epsilon\mathcal{G}(X)) - f^k(X)}{\epsilon}\right) \left.\frac{\partial}{\partial x^k}\right|_X\right) \\
&= \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \frac{\mathcal{G}(X + \epsilon\mathcal{F}(X)) - \mathcal{G}(X)}{\epsilon} - \lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}} \frac{\mathcal{F}(X + \epsilon\mathcal{G}(X)) - \mathcal{F}(X)}{\epsilon} \\
&= \frac{\partial \mathcal{G}}{\partial X}(\mathcal{F}(X)) - \frac{\partial \mathcal{F}}{\partial X}(\mathcal{G}(X)).
\end{aligned} \tag{38}$$

This concludes the proof. □

Theorem 3.1 shows that the Lie product of matrix functions can be calculated by taking limit for given matrix functions without any vectorization as done in [1, 11, 10]. This means that reachability and controllability can be discussed directly in the world of matrix functions. Furthermore, the limit operation is easier to handle than the elementwise calculation. Thus the use of the Lie product of matrix functions with the calculation rule provided by Theorem 3.1 simplifies the reachability and controllability analysis for the systems described by (2).

4 Applications to Reachability Analysis

Here we apply the proposed method to analyze a controlled quantum system and controlled rigid body system to verify its effectiveness.

Analysis of a Controlled Quantum System

Consider a controlled quantum system described by the following Lindblad master equation:

$$\begin{aligned}\frac{\mathrm{d}X}{\mathrm{d}t} &= LXL^* - \frac{1}{2}L^*LX - \frac{1}{2}L^*LX - \mathrm{i}u[H, X] \\ &=: \mathcal{F}[L](X) + \mathcal{G}[H](X)u. \end{aligned} \tag{39}$$

Here we investigate properties of this system using our proposed method. Using Theorem 3.1 and Proposition 2.1, $[\mathcal{F}[L], \mathcal{G}[H]]_{\mathcal{L}}(X)$ is calculated as

$$\begin{aligned}[\mathcal{F}[L], \mathcal{G}[H]]_{\mathcal{L}}(X) = &-\mathrm{i}[H, L]XL^* - \mathrm{i}LX[H, L^*] + \frac{\mathrm{i}}{2}[H, L^*]LX \\ &+ \frac{\mathrm{i}}{2}L^*[H, L]X + \frac{\mathrm{i}}{2}X[H, L^*]L + \frac{\mathrm{i}}{2}XL^*[H, L]. \end{aligned} \tag{40}$$

Assume the following relation holds for a real number β:

$$-\mathrm{i}[H, L] = \beta L. \tag{41}$$

Substituting this equation into (40), we have

$$[\mathcal{F}[L], \mathcal{G}[H]]_{\mathcal{L}}(X) = 2\beta\mathcal{F}[L](X), \tag{42}$$

and thus $\mathscr{C}(X) = \mathrm{span}\{\mathcal{F}[L](X), \mathcal{G}[H](X)\}$. This shows that under the condition (41), local change of quantum system is quite limited.

Note that we do not use any variable transformation in the above analysis unlike the previous works [1, 11, 10].

Analysis of a Controlled Rigid Body System

Attitude of a rigid body is expressed by a 3×3 orthogonal matrix R. Let SO(3) be the set of all orthogonal matrices. Time evolution of the attitude of a rigid body is described by the following equation:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \sum_{i=1}^{3} B_i R \omega_i, \tag{43}$$

where $B_i, i = 1, 2, 3$ is defined as

$$B_1 := E^3_{23} - E^3_{32}, \tag{44}$$
$$B_2 := E^3_{31} - E^3_{13}, \tag{45}$$
$$B_3 := E^3_{12} - E^3_{21}, \tag{46}$$

and $\boldsymbol{\omega}(t) = [\omega_1(t), \omega_2(t), \omega_3(t)]^\top \in \mathbb{R}^3$ denotes the angular velocity of the rigid body (with respect to the axes fixed to the rigid body). This equation can be seen as the following differential equation defined on $\mathbf{M}_3(\mathbb{R})$ with initial condition $X(0) \in \mathrm{SO}(3)$:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \sum_{i=1}^{3} \mathcal{S}_i(X)\omega_i := \sum_{i=1}^{3} B_i X \omega_i. \tag{47}$$

Assume that $\omega_1 \equiv 1, \omega_3 \equiv 0$ and we can change ω_2 (denoted by u hereafter) directly. In this case (47) becomes

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mathcal{S}_1(X) + \mathcal{S}_2(X)u. \tag{48}$$

We investigate the controllability of the rigid body system described by (48) with our proposed method.

Using Theorem 3.1 and Proposition 2.1, $[\mathcal{S}_1, \mathcal{S}_2]_{\mathcal{L}}(X)$ is calculated as

$$\begin{aligned}
[\mathcal{S}_1, \mathcal{S}_2]_{\mathcal{L}}(X) &= \frac{\partial \mathcal{S}_2}{\partial X}(\mathcal{S}_1(X)) - \frac{\partial \mathcal{S}_1}{\partial X}(\mathcal{S}_2(X)) \\
&= B_2 B_1 X - B_1 B_2 X \\
&= B_3 X \\
&= \mathcal{S}_3(X).
\end{aligned} \tag{49}$$

With similar calculations, we get $[\mathcal{S}_2, \mathcal{S}_3]_{\mathcal{L}}(X) = \mathcal{S}_1(X)$ and $[\mathcal{S}_3, \mathcal{S}_1]_{\mathcal{L}}(X) = \mathcal{S}_2(X)$. Thus, $\mathscr{C}(X) = \mathrm{span}\{\mathcal{S}_1(X), \mathcal{S}_2(X), \mathcal{S}_3(X)\}$ holds for any $X \in \mathbf{M}_3(\mathbb{R})$. We can confirm the linear independence of $\mathcal{S}_1(X), \mathcal{S}_2(X)$ and $\mathcal{S}_3(X)$ for $X \in \mathrm{SO}(3)$, thus $\dim \mathscr{C}(X) = 3 = \dim \mathrm{SO}(3)$ on SO(3). By Proposition 1 of [6] we can conclude that the system is controllable, i.e., we can realize any target attitude starting from any initial attitude. This result coincides with the result obtained by different approach [4].

5 Conclusion

In this report we defined the Lie product of matrix functions as a new tool for differential geometric approach to a class of systems described by a matrix-valued differential equation, and provided an efficient calculation rule for the computation. Application of the Lie product of matrix functions with the calculation rule can simplify reachability and controllability analysis for the systems. Specifically, it is expected to be a effective tool to analyze quantum control problems. The effectiveness of the proposed method was verified through examples.

Acknowledgments: This work has been supported in part by Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science, under Grant No. 17656137 and 19560436.

References

[1] C. Altafini. Controllability properties for finite dimensional quantum Markovian master equations. *J. Math. Phys.*, 44(6):2357–2371, 2003.

[2] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, Berlin, 3rd edition, 2004.

[3] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. American Mathematical Society, Providence, R.I., 2001.

[4] S. Jurdjevic and H. J. Sussmann. Control systems on Lie groups. *J. Differential Equation*, 12:313–329, 1972.

[5] G. Lindblad. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.*, 48:119–130, 1976.

[6] C. Lobry. Controllability of nonlinear systems on compact manifolds. *SIAM J. Control*, 12(1):1–4, 1974.

[7] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.

[8] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems*. Springer-Verlag, New York, 1st edition, 1990.

[9] T. Sasaki, S. Hara, and K. Tsumura. Local reachability analysis for controlled quantum dynamics in complex matrix form. In *Proc. 17th Int. Symposium on Mathematical Theory of Networks and Systems*, pages 2001–2008, (Kyoto, Japan), 2006.

[10] N. Yamamoto, K. Tsumura, and S. Hara. Vector representation of stochastic quantum dynamics and its local reachability and observability. Technical Report METR 2004-14, Department of Mathematical Informatics, The University of Tokyo, March 2004.

[11] N. Yamamoto, K. Tsumura, and S. Hara. Local reachability and local observability of controlled quantum dynamics. *Trans. SICE*, 40(11):1078–1087, 2004 (in Japanese).

A Derivation of (37)

$\sum_{j=1}^{N}(\partial g^k/\partial x^j)(X)f^j(X)$ is the abbreviated notation for

$$\sum_{j=1}^{N}\frac{\partial(g^k\circ\varphi^{-1})}{\partial x^j}(\varphi^1(X),...,\varphi^N(X))f^j(X). \tag{50}$$

This is equivalent to the directional derivative of $g^k\circ\varphi^{-1}:\mathbb{R}^N\to\mathbb{R}$ with increment $(f^1(X),...,f^N(X))$ at the point $(\varphi^1(X),...,\varphi^N(X))$, that is,

$$\lim_{\substack{\epsilon\in\mathbb{R}\\ \epsilon\to 0}}\frac{1}{\epsilon}\Big(g^k\circ\varphi^{-1}(\varphi^1(X)+\epsilon f^1(X),...,\varphi^N(X)+\epsilon f^N(X)) \\ -g^k\circ\varphi^{-1}(\varphi^1(X),...,\varphi^N(X))\Big). \tag{51}$$

The following relation is obvious:

$$g^k\circ\varphi^{-1}(\varphi^1(X),...,\varphi^N(X))=g^k(X). \tag{52}$$

In addition, we have

$$\begin{aligned}\varphi^j(X)+\epsilon f^j(X)&=\varphi^j(X)+\epsilon\varphi^j(\mathcal{F}(X))\\ &=\varphi^j(X+\epsilon\mathcal{F}(X)),\quad j=1,...,N\end{aligned} \tag{53}$$

due to the definition of f and the linearity of φ^j. Thus, the following relation holds:

$$\begin{aligned}&g^k\circ\varphi^{-1}(\varphi^1(X)+\epsilon f^1(X),...,\varphi^N(X)+\epsilon f^N(X))\\ &\quad=\ g^k\circ\varphi^{-1}(\varphi^1(X+\epsilon\mathcal{F}(X)),...,\varphi^N(X+\epsilon\mathcal{F}(X)))\\ &\quad=\ g^k(X+\epsilon\mathcal{F}(X)).\end{aligned} \tag{54}$$

Substituting (54) and (52) to (51), we have (37).

B Transformation Rule from Ito Type SDE to Stratonovich Type SDE

It is beneficial to transform an Ito type stochastic differential equation (SDE) to the equivalent Stratonovich type SDE for analyzing the local reachability by calculating the Lie products of matrix functions, because the Stratonovich type SDE has the property of the ordinary chain rule formulas while the Ito type SDE does not [2]. In this appendix, we provide a transformation rule of Ito type SDE to the equivalent Stratonovich type SDE.

Consider the following matrix-valued stochastic differential equation:

$$\begin{aligned} \mathrm{d}X &= \mathcal{F}(X)\mathrm{d}t + \sum_{i=1}^{m} \mathcal{G}_i(X)u_i\mathrm{d}t + \sum_{k=1}^{p} \mathcal{H}_k(X)\mathrm{d}w_k, \\ X(0) &= X_0 \in \mathbf{X}_n. \end{aligned} \tag{55}$$

Here $\mathcal{F}, \mathcal{G}_i, \mathcal{H}_k, i = 1, ..., m, k = 1, ..., p$ are C^∞ matrix functions from $\mathbf{X}_n$ to $\mathbf{X}_n$,$\mathrm{d}w_k(t), k = 1, ..., p$ are standard Wiener increments satisfying $E[\mathrm{d}w_k] = 0$ and $E[\mathrm{d}w_k(t)\mathrm{d}w_l(t)] = \delta_{kl}\mathrm{d}t(\delta_{kl}$ is the Kronecker's delta), and $u_i(t) \in \mathbb{R}, i = 1, ..., m$ are control inputs.

The following result provides the direct transformation rule of Ito type SDE to the equivalent Stratonovich type SDE.

Lemma B.1 *Ito type SDE (55) is equivalent to the following Stratonovich type SDE:*

$$\begin{aligned} \mathrm{d}X &= \left(\mathcal{F}(X) - \frac{1}{2}\sum_{k=1}^{p} \frac{\partial \mathcal{H}_k}{\partial X}(\mathcal{H}_k(X)) \right) \mathrm{d}t + \sum_{i=1}^{m} \mathcal{G}_i(X)u_i\mathrm{d}t \\ &\quad + \sum_{k=1}^{p} \mathcal{H}_k(X) \circ \mathrm{d}w_k, \\ X(0) &= X_0 \in \mathbf{X}_n. \end{aligned} \tag{56}$$

Proof *Equation (55) stands for the following set of N real scalar valued Ito type SDEs:*

$$\begin{aligned} \mathrm{d}x^j &= f^j(X)\mathrm{d}t + \sum_{i=1}^{m} g_i^j(X)u_i\mathrm{d}t + \sum_{k=1}^{p} h_k^j(X)dw_k, \\ &= f^j(X)\mathrm{d}t + \sum_{i=1}^{m} g_i^j(X)u_i\mathrm{d}t + \sum_{k=1}^{p} h_k^j(X)\mathrm{d}w_k, \\ x^j(0) &= x_0^j, \quad j = 1, ..., N. \end{aligned} \tag{57}$$

We can check it by parameterizing the equation (55) according to (6) or (7), respectively. They are equivalent to the following set of N real scalar valued

Stratonovich type SDEs [2]:

$$\begin{aligned}
\mathrm{d}x^j &= \left(f^j(X) - \frac{1}{2}\sum_{k=1}^{p}\sum_{l=1}^{N} \frac{\partial h_k^j}{\partial x^l}(X) h_k^l(X) \right)\mathrm{d}t + \sum_{i=1}^{m} g_i^j(X) u_i \mathrm{d}t \\
&\quad + \sum_{k=1}^{p} h_k^j(X) \circ \mathrm{d}w_k \\
x^j(0) &= x_0^j, \quad j = 1, ..., N. \qquad (58)
\end{aligned}$$

The remainder of the proof is same as that of Lemma 3.1. The key point is rewriting the right hand side of the equation (58) with limit operation. □